

# A Smart Way to Manage OSS Compliance with Yocto+SPDX

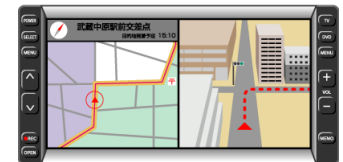
Jul 13th, 2016

Lei Maohui, Fujitsu

leimaohui@cn.fujitsu.com



- Working for Fujitsu from 2011
- 3 years experience in Yocto related development
- In-House Embedded Linux Distributor of Fujitsu
- Our Distribution includes LTSI Kernel and is built with Yocto Project
- Our Distribution is used for
  - IVI
  - Server System Controller
  - Storage System
  - Network Equipment
  - Printer
  - etc.

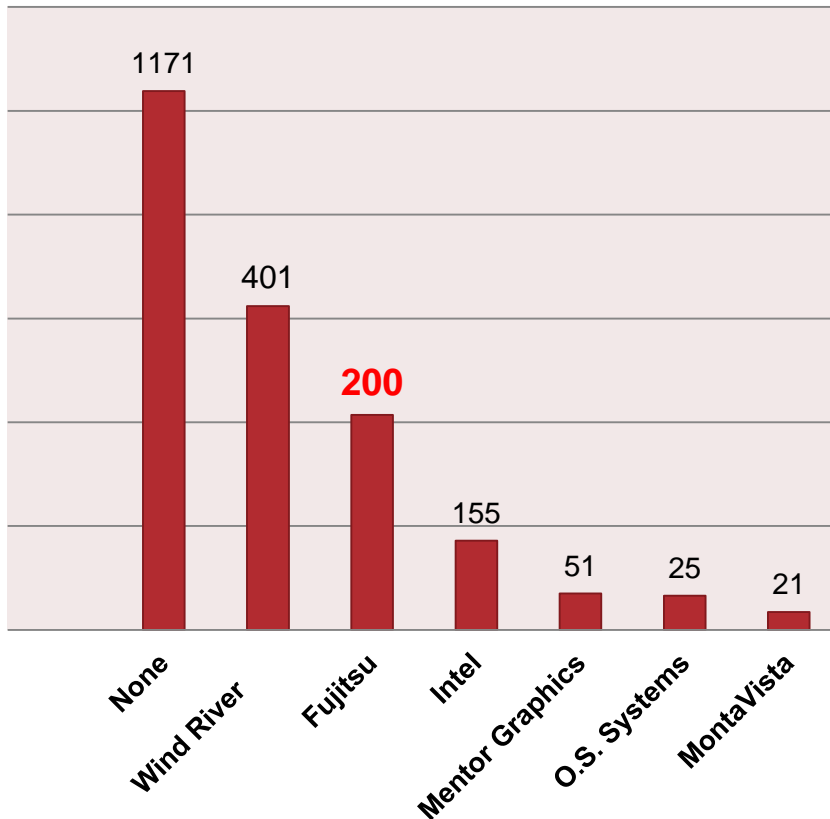


# Our contributions to Yocto community

■ Data comes from meta-openembedded.git ( 2015-01-01 ~ 2016-06-30)

Top changeset contributors by employer

■ commits



	Developer	Changesets
1	Martin Jansa	235 (10.6%)
2	Andreas Müller	222 (10.0%)
3	Li Xin (Fujitsu)	100 (4.5%)
4	Roy Li	84 (3.8%)
5	Alexander Kanavin	69 (3.1%)
6	Yi Zhao	64 (2.9%)
7	Kai Kang	53 (2.4%)
8	Andre McCurdy	41 (1.8%)
9	Jackie Huang	39 (1.8%)
10	Bian Naimeng (Fujitsu)	38 (1.7%)
11	Khem Raj	36 (1.6%)
12	Maohui Lei (Fujitsu)	32 (1.4%)
13	Paul Eggleton	31 (1.4%)
14	Joe MacDonald	29 (1.3%)
15	Tim Orling	29 (1.3%)

Developers with the most changesets

## Introduction of SPDX

- In your company
- What is SPDX
- Who are working for SPDX
- The status of SPDX specification

## Yocto+SPDX

- Current state
- Current problems of Yocto+SPDX

## What we have done

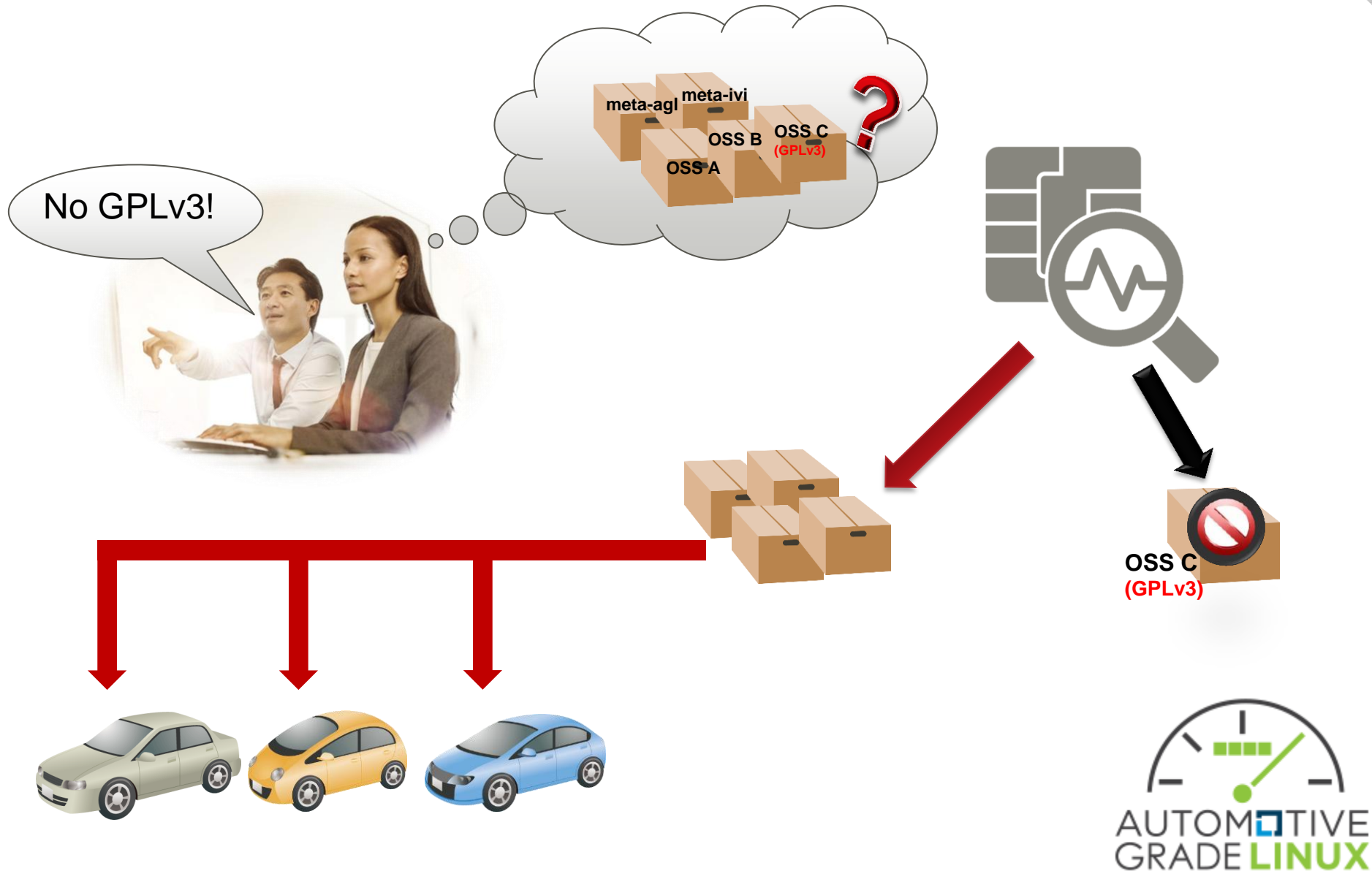
- Aim to make Yocto+SPDX support SPDX 1.2
- Aim to make Yocto+SPDX support SPDX 2.0
  - Current SPDX create tools
  - Our contributions
  - Before and after

# Introduction of SPDX

- In your company
- What is SPDX
- Who are working for SPDX
- The status of SPDX specification



# In your company



## What is SPDX

- The full name of SPDX is **S**oftware **P**ackage **D**ata **E**xchange, which is a standard format for communicating the components, licenses and copyrights associated with a software package.

## Vision of SPDX

- To help reduce redundant work in determining software license information and facilitate compliance.

If you are:

- Developers of open source projects
- Developers of software that includes a Linux distro

You can get the following information of OSS for your users or customers:

- Licenses
- Copyrights

**SPDX** will be a good solution, if a SPDX implementation can generate SPDX file including license information automatically.



# Who are working for SPDX

## primary responsibility

- Drafts the specification
- develops documentation templates, samples and tools.

## Delivered

- SPDX Spec (2.0,1.2,1.1,1.0)
- Tool (fossology)
- Spreadsheet Template

## Recent

- SPDX Specification 2.1
- tool for 2.1

## Primary responsibility

- Supports and provides recommendations to the SPDX working groups regarding licensing issues.
- Maintains the [SPDX License List](#)
- Promotes the SPDX specification to the legal community at-large

## Delivered

- License Expression Syntax
- License Inclusion Guidelines (Background)
- Dealing with Public Domain within SPDX Files

## Recent

- Primary focus getting all the licenses into GitHub
- New licenses

Technical Team

General Meetings

Legal Team

Business Team

## Primary responsibility

- Launch activities for new versions of the SPDX specification.
- Outreach
- Participation in events;
- The SPDX website

## Delivered

- Launch for 1.0 and 1.1
- Process for Adding to License List (Draft)
- SPDX Vision & Mission Discussion Document
- SPDX Vision & Mission Statements (Final Draft)

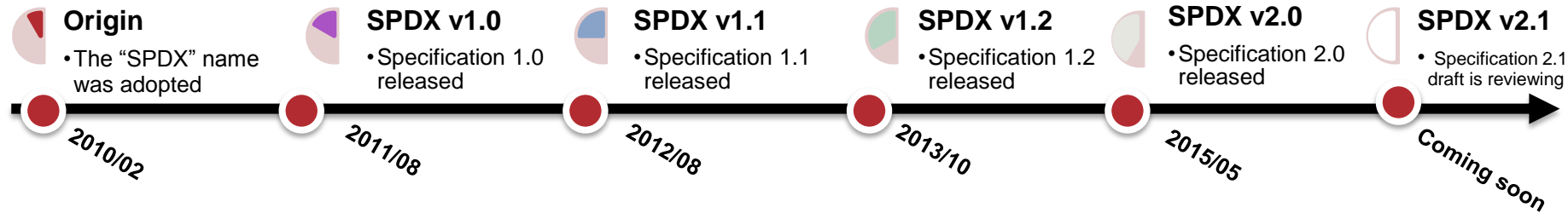
## Recent

- The SPDX website

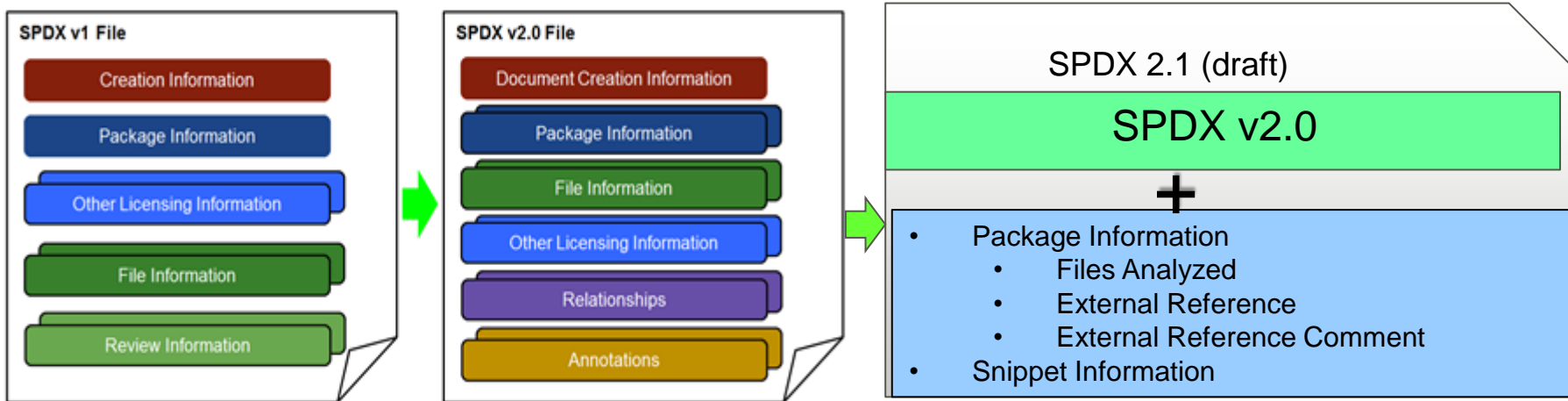
- Obtain details from <http://spdx.org/participate>

# The status of SPDX Specification

## History



## New features in SPDX v2.x



## Obtain details from

- <https://spdx.org/about-spdx/what-is-spdx>
- [http://wiki.spdx.org/view/Technical\\_Team/SPDX\\_Specification\\_Versions](http://wiki.spdx.org/view/Technical_Team/SPDX_Specification_Versions)
- [http://spdx.org/sites/spdx/files/publications/SPDX\\_2.0\\_Collab\\_2015.pdf](http://spdx.org/sites/spdx/files/publications/SPDX_2.0_Collab_2015.pdf)

# Yocto+SPDX

- Current state
- Current problems of Yocto+SPDX



## Status

### History

- Yocto+SPDX was supported from yocto 1.5.

### SPDX Specification

- Yocto+SPDX supports SPDX v1.1 specification.

### SPDX Implementation

- Yocto+SPDX generates spdx files by using fossology-spdx server.

## Activity of Yocto+SPDX

- There are almost no improvements in spdx module.

```
$ git log --pretty=format:"%ad %s" meta/classes/spdx.bbclass
Thu Nov 5 17:48:18 2015 +0200 bbclass: fix spelling mistakes
Thu Nov 13 15:49:52 2014 +0100 spdx.bbclass: improved error handling and code cleanup
Mon Oct 20 16:09:15 2014 +0200 spdx.bbclass: improved stability, fixed SPDX compliance issues. Changes are reflected in licenses.conf.
Tue Sep 23 17:48:12 2014 +0800 spdx.bbclass: Add SPDX-specific source tree variable.
Sun Sep 1 08:52:40 2013 +0100 meta: Don't use deprecated bitbake API
Fri Aug 23 14:40:35 2013 -0700 SPDX:real-time license scanning and SPDX output.
```

# Current problems of Yocto+SPDX

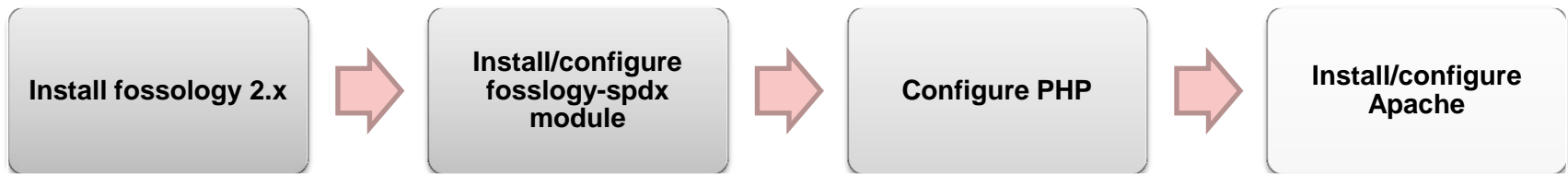
## Only support SPDX v1.1

- Even SPDX v1.1, Yocto+SPDX doesn't support well.

Section	Fields	Mandatory	Yocto+SPDX
Creation Information	Creator	Yes	NO
Package Information	Package Download Location	Yes	NO
	All Licenses Information from Files	Yes	NO
	Declared License	Yes	NO

## Complex

- Complex to build a Yocto+SPDX environment



## Poor performance

- Create a sdx file will spend too much time



# What we have done

- Make Yocto+SPDX support SPDX 1.2
- Make Yocto+SPDX support SPDX 2.0
  - Current SPDX create tools
  - Our contributions
  - Before and after

# Make Yocto+SPDX support SPDX 1.2 (1/2)

## Deviations from SPDX 1.2 specification

No.	Section	Field	Mandatory	Already in Yocto+SPDX	Compliance With SPDX-1.2	Need to be fixed
1	SPDX Document Information	SPDX Version	Yes	Yes	No	Yes
2		Data License	Yes	Yes	Yes	No
3		Document Comment	No	Yes	Yes	No
4	Creation Information	Creator	Yes	Yes	No	Yes
5		Created	Yes	Yes	Yes	No
6		Creator Comment	No	Yes	Yes	No
7		License List Version	No	No	Unkown	No
8	Package Information	Package Name	Yes	Yes	Yes	No
9		Package Version	No	Yes	Yes	No
10		Package File Name	No	Yes	Yes	No
11		Package Supplier	No	Yes	Yes	No
12		Package Originator	No	Yes	Yes	No
13		Package Download Location	Yes	Yes	No	Yes
14		Package Verification Code	Yes	Yes	Yes	No
15		Package Checksum	No	Yes	Yes	No
16		Package Home Page	No	No	Unkown	No
17		Source Information	No	No	Unkown	No
18		Concluded License	Yes	Yes	Yes	No
19		All Licenses Information from Files	Yes	Yes	No	Yes
20		Declared License	Yes	Yes	No	Yes
21		Comments on License	No	No	Unkown	No
22		Copyright Text	Yes	Yes	Yes	No
23		Package Summary Description	No	Yes	No	Yes
24		Package Detailed Description	No	Yes	Yes	No
25	Other Licensing Information Detected	Identifier Assigned	Conditional	No	No	Yes
26		Extracted Text	Conditional	No	No	Yes
27		License Name	Conditional	No	No	Yes
28		License Cross Reference	No	No	Unkown	No
29		License Comment	No	No	Unkown	No
30	File Information	File Name	Yes	Yes	Yes	No
31		File Type	No	Yes	Yes	No
32		File Checksum	Yes	Yes	Yes	No
33		Concluded License	Yes	Yes	Yes	No
34		License Information in File	Yes	Yes	Yes	No
35		Comments on License	No	No	Unkown	No
36		Copyright Text	Yes	Yes	Yes	No
37		Artifact of Project Name	No	No	Unkown	No
38		Artifact of Project Homepage	No	No	Unkown	No
39		Artifact of Project Uniform Resource Identifier	No	No	Unkown	No
40		File Comment	No	No	Unkown	No
41		File Notice	No	No	Unkown	No
42		File Contributor	No	No	Unkown	No
43		File Dependencies	No	No	Unkown	No
44	Review Information	Reviewer	No	No	Yes	No
45		Review Date	Conditional	No	Unkown	No
46		Review Comment	No	No	Yes	No

- Make Yocto+SPDX be compliant with SPDX-1.2 specification

发件人: ☐ Lei, Maohui  
收件人: ☐ openembedded-core@lists.openembedded.org  
抄送: ☒ Lei, Maohui  
主题: [PATCH v4 0/3] These patches aim to make the spdx file be compliant with the SPDX 1.2 Specification.

发送时间: 2015/6/10 (周二) 18:09

Those patches aim to make the spdx file be compliant with the SPDX 1.2 Specification.

If you want to use this feature, you need to do:

1. Make sure your fossology+spdx server works well. You can get spdx file with the following command.  
curl <http://127.0.0.1/repo/> --no-proxy 127.0.0.1 -k -F mod=spdx\_license\_once -F noCopyright=false  
-F jsonOutput=false -F fullSPDXFlag=true -F [file=@xxx.tar.gz](#) -o xxx.spdx
2. Add the following INHERIT statement and set the SPDX\_MANIFEST\_DIR at the end of your conf/local.conf file found in the Build Directory.  
INHERIT += "spdx"

Lei Maohui (3):

licenses.conf: Modified parameters for new spdx.bbclass  
spdx.bbclass: Create the spdx file which is compliant with SPDX  
1.2 Specification  
spdx: create a directory to save source code

meta/classes/spdx.bbclass | 431 ++++++-----  
meta/conf/licenses.conf | 86 +++  
2 files changed, 184 insertions(+), 333 deletions(-)

--  
1.8.4.2

- **This patch has not been merged into mainline tree. But already been used by some people or companies.**

# Make Yocto+SPDX support SPDX 2.0

## AIMS

- Support SPDX 2.0
- Good performance
- Easy to build a Yocto+SPDX environment

## SPDX Create Tools

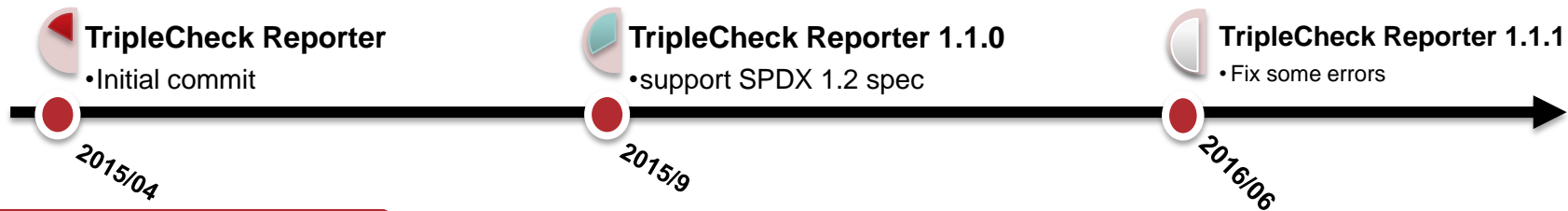
- TripleCheck Reporter
- FOSSology3
- DoSOCSv2



## • What is TripleCheck Reporter

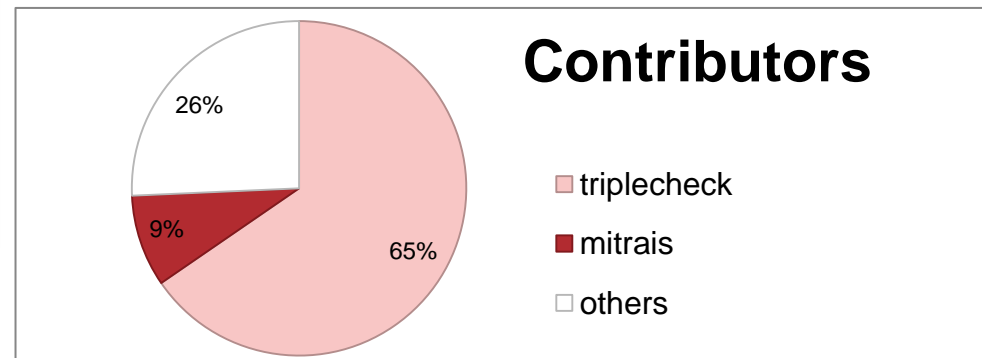
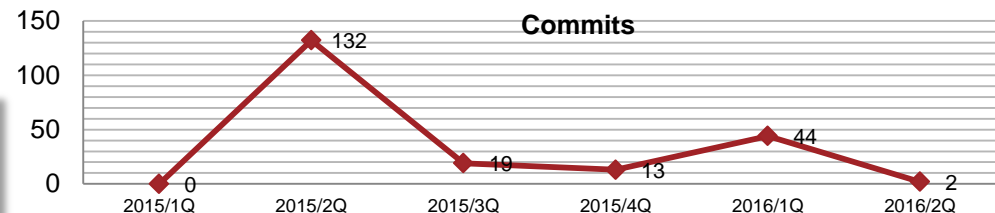
- The TripleCheck reporter is the ideal tool for a quick overlook of the licensing compliance status for a given set of source code files in your desktop computer (Linux, Windows and Mac OS X). ([Website](#))

## • History



## • Project Activity

Item		TripleCheck
Last release		2016/06
Contributors	All Time	5
	Past 12 Months	3
Commits	All Time	224
	Past 12 Months	92
Activity level		Low



- (1) Data comes from OpenHub - [www.openhub.net](http://www.openhub.net).
- (2) Git Repository: <https://github.com/triplecheck/triplecheck.github.io>

## • How to use TripleCheck Reporter

Reports (1)

cpio-2.11

Files (524)

build-aux

src

m4

headers

po

gnu

tests

rmt

am

lib

doc

Makefile.am (C)

THANKS

ChangeLog (N)

/Makefile.in (c)

aclocal.m4

AUTHORS

INSTALL

config.h.in

TODO

configure.ac (C)

COPYING (Not)

/configure (c)

README

ABOUT-NLS

Make.rules (GP)

NEWS (GPL-3.0)

Authorshin

Search files..

cpio-2.11

67,253 lines of code

524 files in total

6 Mb in size

41%

59%

License declared

No license declared

147 files

308 files

License evidence

Declared license

[none]

75% GPL-3.0+ (231 files) [View](#)

10.4% GPL-3.0 (32 files) [View](#)

7.5% Not recognized (23 files) [View](#)

2.6% LGPL-3.0+ (8 files) [View](#)


1.9% Public\_Domain (6 files) [View](#)

1.9% GPL-2.0+ (6 files) [View](#)

0.3% MIT, Public Domain (1 file) [View](#)

```

##-----
## SPDX Document Information
##-----
SPDXVersion: SPDX-1.2
DataLicense: CC0-1.0
##-----
## Creation Information
##-----
Creator: Person: root
Creator: Tool: TripleCheck 1.1.1
Created: 2016-07-10T18:05:33Z
##-----
## Package Information
##-----
PackageName: cpio-2.11
PackageLicenseDeclared: NOASSERTION
##-----
## File Information
##-----
FileName: ./Makefile.am
FileType: OTHER
FileChecksum: SHA1:
3b7a83d30d4ade6e3ad48f3da5f957a9da2078d9
FileChecksum: SHA256:
82c2300ee5d555edea3d038398ab1c13f76ff6931158
d474a5bedae5b4ad2fe3
FileChecksum: MD5:
b7310cf3b97f49a1335fa68fb5b6c93c
FileSize: 1024 bytes (1024 bytes)
FileCopyrightText: <text></text>
LicenseInfoInFile: GPL-3.0+
    
```



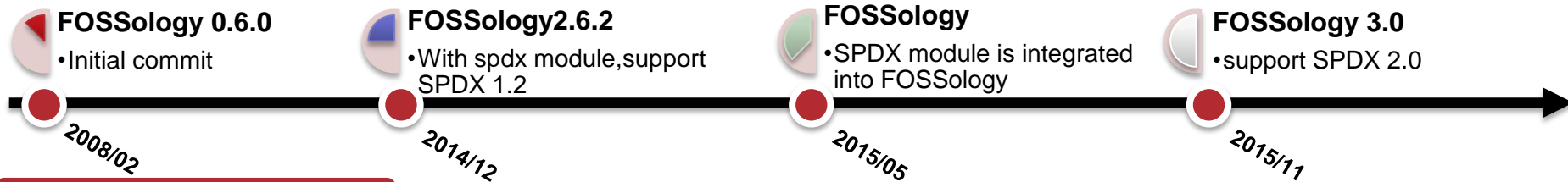
17

Copyright 2016 FUJITSU COMPUTER TECHNOLOGIES LIMITED

## • What is FOSSology

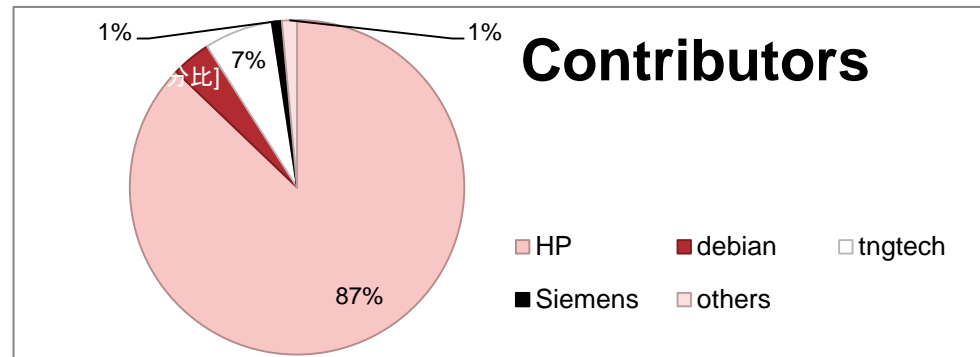
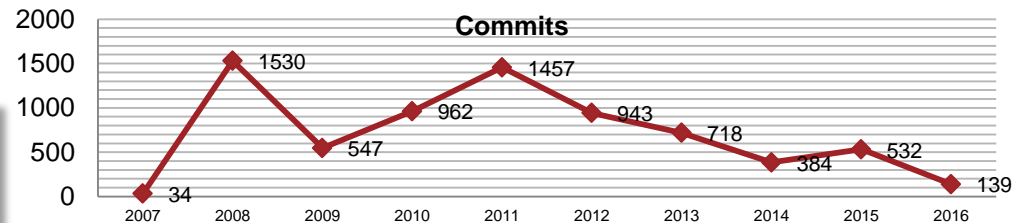
- FOSSology is a open source license compliance software system and toolkit. As a toolkit you can run license, copyright and export control scans from the command line. As a system, a database and web ui are provided to give you a compliance workflow. License, copyright and export scanners are tools available to help with your compliance activities.([Website](#))

## • History



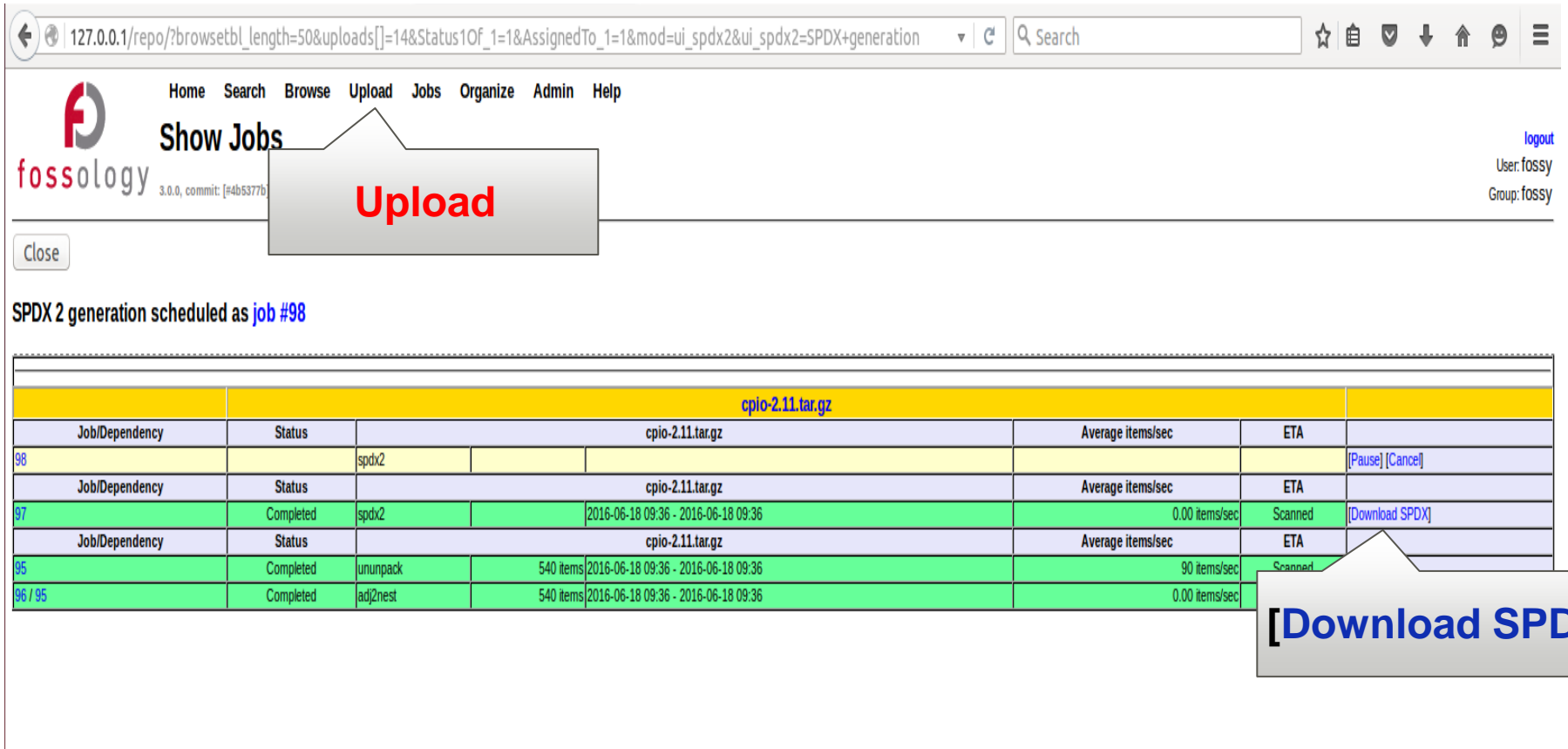
## • Project Activity

Item		FOSSology
Last release		2015-11
Contributors	All Time	45
	Past 12 Months	24
Commits	All Time	7,294
	Past 12 Months	409
Activity level		Moderate



- (1) Data comes from OpenHub - [www.openhub.net](http://www.openhub.net).
- (2) Git Repository: <https://github.com/fossology/fossology>  
<https://github.com/FOSSology-SPDX/fossology-spdx>

- How to use FOSSology3



The screenshot shows the FOSSology3 web interface. At the top, there is a navigation bar with links: Home, Search, Browse, Upload, Jobs, Organize, Admin, Help. The 'Upload' link is highlighted with a red box and the word 'Upload' in red text. Below the navigation bar, there is a 'Show Jobs' button and a 'Close' button. The main content area displays the text 'SPDX 2 generation scheduled as job #98'. Below this, there is a table showing the progress of the SPDX generation process.

Job/Dependency	Status	cpio-2.11.tar.gz	Average items/sec	ETA
98		spdx2		
97	Completed	spdx2	0.00 items/sec	Scanned
95	Completed	ununpack	540 items	2016-06-18 09:36 - 2016-06-18 09:36
96 / 95	Completed	adj2nest	540 items	2016-06-18 09:36 - 2016-06-18 09:36

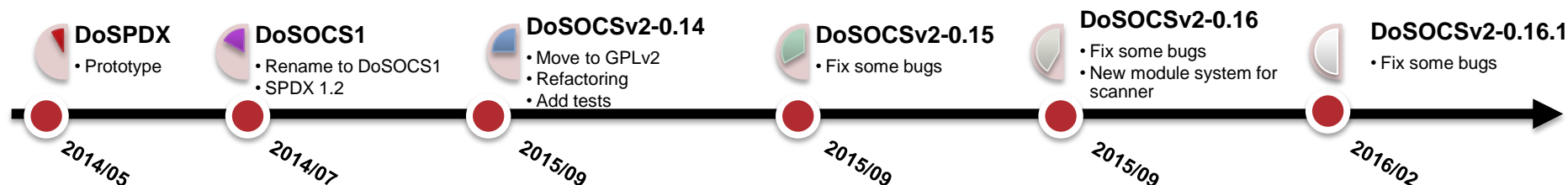
Buttons: [Pause] [Cancel] [Download SPDX]

# Current SPDX create tools - DoSOCSv2 (1/2)

## What is DoSOCSv2

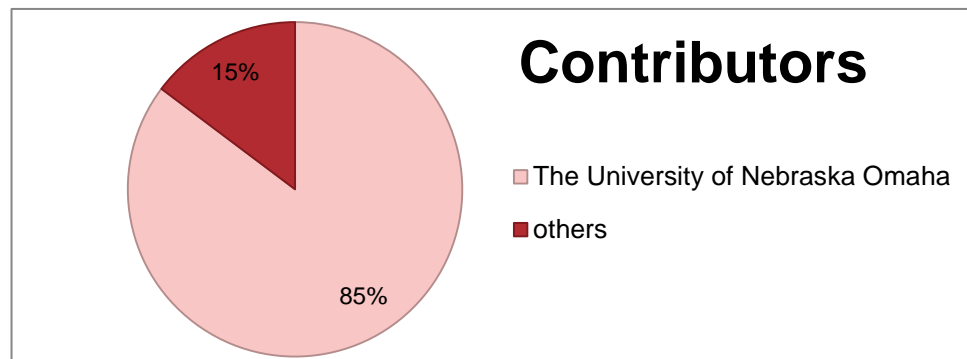
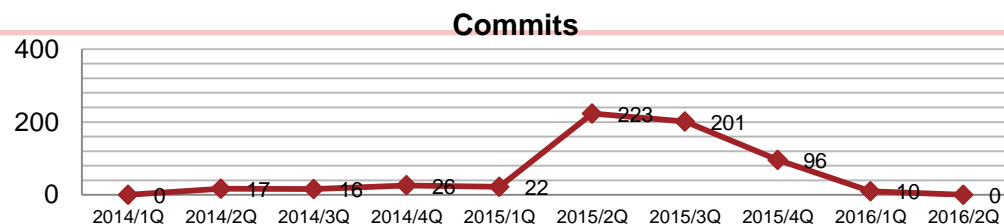
- dosocs2 is a command-line tool for managing SPDX 2.0 documents and data. It can scan source code distributions to produce SPDX information, store that information in a relational database, and extract it in a plain-text format on request. ([Website](#))

## History



## Project Activity

Item		DoSOCSv2
Last Release		2016/02
Contributors	All Time	12
	Past 12 Months	7
Commits	All Time	611
	Past 12 Months	495
Activity Level		Moderate



- (1) Data comes from OpenHub - [www.openhub.net](http://www.openhub.net).
- (2) Git Repository: <https://github.com/DoSOCSv2/DoSOCSv2>

## How to use DoSOCSv2

```
$ dosocs2 oneshot cpio-2.11
dosocs2: cpio-2.11: package_id: 1
dosocs2: running nomos on package 1
cccccpio-2.11: document_id: 1
```

### **SPDXVersion: SPDX-2.0**

```
DataLicense: CC0-1.0
DocumentNamespace: sqlite:///home/leimh/.config/dosocs2/dosocs2.sqlite3/cpio-2.11-fe30375e-3a43-4d1e-9962-eb24f2dbe8bf
DocumentName: cpio-2.11
SPDXID: SPDXRef-DOCUMENT
DocumentComment: <text></text>
```

#### ## External Document References

#### ## Creation Information

```
Creator: Tool: dosocs2-0.16.1
Created: 2016-07-09T23:18:52Z
CreatorComment: <text></text>
```

### **LicenseListVersion: 2.2**

#### ## Document Annotations

#### ## Document Relationships

```
Relationship: SPDXRef-DOCUMENT DESCRIBES SPDXRef-package-cpio_2_11-f6eb-4fa85311
Relationship: SPDXRef-DOCUMENT DESCRIBES SPDXRef-file-ABOUT_NLS-b502-579bb6d1
Relationship: SPDXRef-DOCUMENT DESCRIBES SPDXRef-file-AUTHORS-2cd7-1fb19a33
Relationship: SPDXRef-DOCUMENT DESCRIBES SPDXRef-file-COPYING-8427-1a9a3562
Relationship: SPDXRef-DOCUMENT DESCRIBES SPDXRef-file-ChangeLog-6f23-76c9a0d2
.....
```

# DoSOCSv2 is best for Yocto

Item		TripleCheck Reporter	FOSSology3	DoSOCSv2
Last Release		1.1.1	3.0	v0.16.0
License		AGPLv3	GPLv2	GPLv2
Support SPDX version		1.2	<b>2.0</b>	<b>2.0</b>
Scanners		N/A	Nomos, Monk, Ninka	Nomos
Supported Platform	Linux	√	√	√
	Others (Windows/OS X)	√		
Interface adapt to Yocto			√ (Partial support)	√
Graphical user interface		√	√	
Project Activity ( <a href="http://www.openhub.net">http://www.openhub.net</a> )		Low	Moderate	Moderate
Scan time		<b>Short</b>	Long	Middle
Scan unpacked sources		√		√
Build environment complexity		<b>Easy</b>	complex	<b>Easy</b>

## Yocto+SPDX switch to DoSOCSv2

- No need to pack & unpack source code
- The scan results and other collected metadata are saved in the database so that subsequent document generations will be much faster.

发件人: ☐ Lei, Maohui

收件人: ☐ openembedded-core@lists.openembedded.org

发送时间: 2016/6/27 (周一) 8:12

抄送: ☐ Lei, Maohui

主题: [OE-core][PATCH] To make yocto-spx support spdx2.0 SPEC

There are some problems in spdx module(spx.bbclass).

1. The newest version of spdx specification is 2.0. But even spdx 1.1, yocto+SPDX can't support well.
2. It is complex to build a Yocto+SPDX environment.
3. Creating a spdx file spends too much time, especially for large software.

To improve spdx module ,I change the spdx create tool from fossology to dosocs2.

With this patch:

1. Also gets license informations by scanner from fossology.

1. Can support SPDX2.0 SPEC.

2. Because dosocs2 can work on directories, so there is no necessary to pack source code before do\_spx. It can save time for large software.

Lei Maohui (1):

To Make yocto-spx support spdx v2.0 specification.

```
meta/classes/spx.bbclass | 488 ++++++-----
meta/conf/licenses.conf   | 67 +-----
2 files changed, 184 insertions(+), 371 deletions(-)
```

--

1.9.1

---

```
meta/classes/spx.bbclass | 488 ++++++-----
meta/conf/licenses.conf   | 67 +-----
2 files changed, 184 insertions(+), 371 deletions(-)
```

```
diff --git a/meta/classes/spx.bbclass b/meta/classes/spx.bbclass index 0c92765..892199d 100644
```

```
--- a/meta/classes/spx.bbclass
```

```
+++ b/meta/classes/spx.bbclass
```

```
@@ -1,12 +1,9 @@
```

```
# This class integrates real-time license scanning, generation of SPDX standard # output and verifying license info during the building process.
```

```
-# It is a combination of efforts from the OE-Core, SPDX and Fossology projects.
```

```
+# It is a combination of efforts from the OE-Core, SPDX and DoSOCSv2 projects.
```

```
#
```

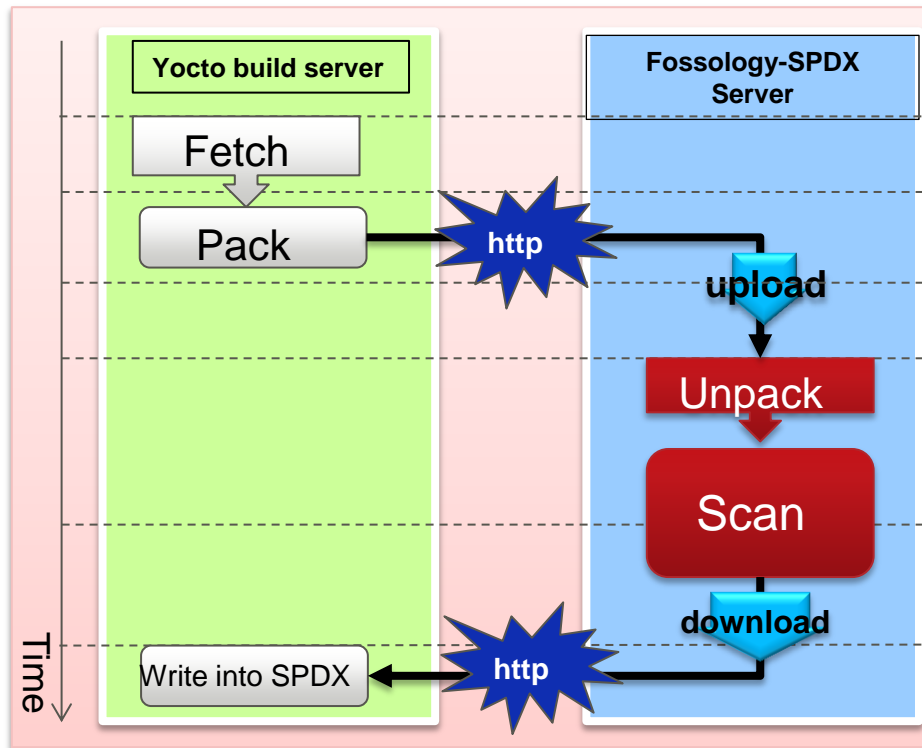
```
-# For more information on FOSSology:
```

```
-# http://www.fossology.org
```

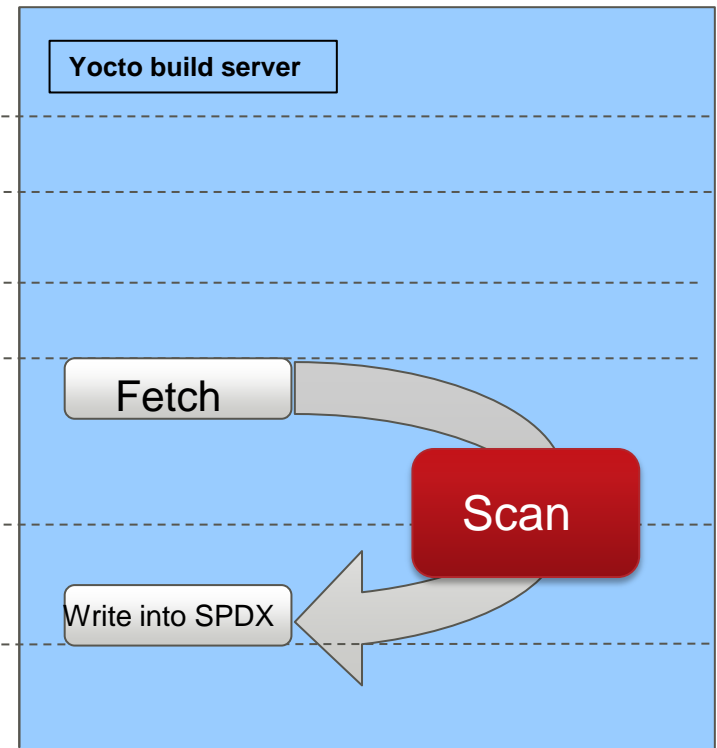
```
..
```

# Before and after (1/3)

## Before



## After



# Before and after (2/3)

- With our patch, Yocto+SPDX has better performance after first time.



# Before and after (3/3)

Item		Before	After
SPDX version		SPDX 1.1	SPDX 2.0
SPDX create tool		fossology-spdx	dosocs2
Scanner		nomos	nomos
LicenseListVersion		1.19	2.2
Performance (e.g. glibc- 2.24)	First time	75min	44min
	Second time	77min	6min

- Current deviations from SPDX 2.0 specification
  - Exactly one package per document is required. (SPDX 2.0 allows zero or more packages per document.)
  - Files in a document can only exist within a package. (SPDX 2.0 allows files to exist outside of a package.)
  - Checksums are always assumed to be SHA-1. (SPDX 2.0 permits SHA-1, SHA-256, and MD5)
  - A file may be an artifact of only one project.
  - License expression syntax is not parsed; license expressions are interpreted as license names that are not on the SPDX license list.
  - Deprecated fields from SPDX 1.2 (reviewer info and file dependencies) are not supported.

# Need to improve (2/2)

## Only one scanner

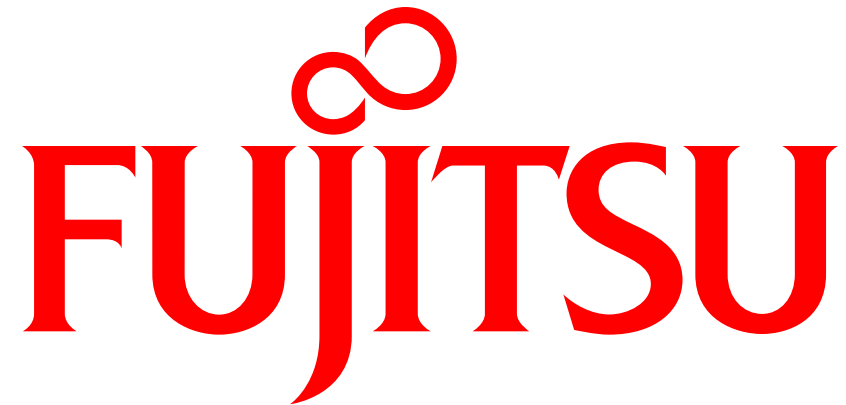
- By default, dosocs2 only supports nomos as scanner, less than fossology3.
- Scanners provided by fossology3

Scanner	Overview	History
Nomos	<ul style="list-style-type: none"><li>• License identification is done using short phrases (regular expressions) and heuristics. The heuristics for detecting phrases must be found in (or out of) proximity to another phrase or phrases.</li><li>• This scanner currently recognizes more than 659 licenses</li></ul>	<ul style="list-style-type: none"><li>• This is one of the original scanners used by HP that was the foundation for FOSSology when it was open sourced in 2007, and has been maintained and enhanced through the years</li></ul>
Ninka	<ul style="list-style-type: none"><li>• Ninka is sentence-based, and provides a simple way to identify open source licenses in a source code file.</li><li>• Ninka was designed to be lightweight, fast, and if it isn't sure about the license, not to guess.</li></ul>	<ul style="list-style-type: none"><li>• This scanner was developed by a team of researchers studying automatic license detection in 2010</li><li>• Ninka has been incorporated into the FOSSology Project as part of the 3.0 release in 2015. It is used in other open source projects as well.</li></ul>
Monk	<ul style="list-style-type: none"><li>• Monk looks for complete licenses (as defined in the license database) and reports the percentage of match to that reference version. It is useful with license highlighting, as it allows you to see exactly what was added or removed from a license.</li><li>• Text similarity is based on the Jaccard index</li></ul>	<ul style="list-style-type: none"><li>• This scanner was contributed by Siemens and TNGtech to the FOSSology project.</li><li>• It was made available as part of the 2.6 release in 2014.</li></ul>

Scanners information comes from: <http://events.linuxfoundation.org/sites/events/files/fossology-overview-20151109.1.pdf>

- Introduce spdx
- Introduce the problems of Yocto+SPDX.
- What we have done for Yocto+SPDX.
  - Make Yocto+SPDX switch to DoSOCSv2.
- Need to improve for our contribution.

# Any Questions?



shaping tomorrow with you